EXHIBIT C

<div align="center">

SUSMAN GODFREY L.L.P.

A REGISTERED LIMITED LIABILITY PARTNERSHIP

ONE MANHATTAN WEST

NEW YORK, NEW YORK 10001-8602

(212) 336-8330

FAX (212) 336-8340

WWW.SUSMANGODFREY.COM

_____

</div>

Suite 5100	Suite 1400	Suite 3000
1000 Louisiana Street	1900 Avenue of the Stars	401 Union Street
Houston, Texas 77002-5096	Los Angeles, California 90067-6029	Seattle, Washington 98101-2683
(713) 651-9366	(310) 789-3100	(206) 516-3880

Zach Savage
Direct Dial (212) 729-2022

E-Mail ZSavage@susmangodfrey.com

April 19, 2024

VIA ELECTRONIC MAIL

Andrew Gass
Andrew.gass@lw.com
505 Montgomery Street, Suite 2000
San Francisco, California 94111-6538

> Re: *The New York Times Company v. Microsoft Corporation, et al.*,
> Case No.: 23-cv-11195-SHS

Counsel:

I write regarding OpenAI's Responses and Objections to Plaintiff's First Set of Requests for Production served on March 25, 2024 (the "R&Os"). Please agree in writing to remedy the deficiencies identified below or offer a time to meet and confer no later than April 26, 2024.

1. Limitation to Inspection of "Pretraining Data" (RFP Nos. 1, 3, 8, 14)

In response to RFPs 1, 3, 8, and 14, each of which requests discrete categories of information and documents, OpenAI states (after a lengthy series of objections) that it will "make available for inspection, pursuant to an inspection protocol to be negotiated between the parties, the pretraining data for models used for ChatGPT that it locates after a reasonable search."

That blanket response is deficient. For starters, OpenAI does not state, as is required by Federal Rule 34(b)(2)(C), whether it is seeking to withhold responsive documents on the basis of its objections (and on which ones). We presume, given the limited scope of documents OpenAI is currently agreeing to produce in response to these requests, that OpenAI is seeking to withhold such documents, but OpenAI should make that position clear. Please amend your responses accordingly.

April 19, 2024
Page 2

Second, assuming that OpenAI is, in fact, seeking to limit its production obligations in response to RFPs 1, 3, 8, and 14 solely to pretraining data, that limitation is improper. These RFPs go far beyond a request for pretraining data. For example:

- RFP No. 1: Requests documents concerning any rules or guidance for choosing what is and is not included in the Training Datasets.
- RFP No. 3: Requests documents concerning the "appropriateness, value, importance, necessity, and/or concerns over or legal implications of" using copyrighted materials for training and other purposes.
- RFP No. 8: Requests documents concerning the form in which Times Content is stored in the AI Models.
- RFP No. 14: Requests the source from which each data sample contained in the Training Datasets was obtained.

None of these requests is likely to be fully encompassed by pretraining data. The above discussion is not intended to be exhaustive, but is simply intended to illustrate the scope of the requests and OpenAI's improper narrowing of them. Please let us know whether you will reconsider your position with respect to RFPs 1, 3, 8, and 14.

Please also confirm, pursuant to RFP No. 1, that OpenAI will produce documents related to the "use of Times Content for any *post-training* activity, including fine tuning." R&Os at 7 (emphasis added).

2. Limiting Definitions of Generative AI Products and Services and AI Models (RFP Nos. 5, 12, 18)

OpenAI's responses to The Times's requests are deficient to the extent that they unilaterally limit and redefine The Times's requests by interpreting "Generative AI Products and Services" to mean only ChatGPT. *See, e.g.*, R&Os at 5, 12, 18. Particularly in light of OpenAI's substantial collaboration with Microsoft, "Generative AI Products and Services" includes CoPilot, Browse with Bing, Bing Chat, APIs, and all other platforms otherwise providing access to ChatGPT. *See* Compl. ¶¶ 16, 66–74, 108–23 (addressing these additional products and services). OpenAI's responses are likewise deficient to the extent that they redefine "AI Models" to mean "only the models used for ChatGPT" which "are more limited in scope." R&Os at 5. Again, in light of the collaboration between Microsoft and OpenAI, "AI Models" includes the models used to train any of Defendants' other AI products. *See* Compl. ¶¶ 66–74 (describing Microsoft's "intimate[] involve[ment] in the training, development, and commercialization of OpenAI's GPT products").

April 19, 2024
Page 3

3. <u>Limitation to "Copyright Shield" Policy Documents (RFP No. 5)</u>

RFP 5 requests documents reflecting policies, procedures or practices concerning the use of intellectual property in AI Models, *including* the "Copyright Shield" policy. In response, after a lengthy list of objections, OpenAI states that it will produce documents "sufficient to explain the 'Copyright Shield' policy." R&Os at 11.

That response is inadequate for several reasons. First, OpenAI again fails to state whether it is seeking to withhold responsive documents on the basis of its objections (and which ones). It appears that OpenAI is seeking to do so, but please amend your responses to confirm as much. Second, OpenAI is agreeing to produce documents that "explain" the Copyright Shield policy. We do not understand what it means to "explain" the Copyright Shield policy, nor is The Times's request so limited. Third, RFP 5 requests documents reflecting policies, procedures or practices concerning the use of intellectual property in AI Models; it is not limited to the Copyright Shield policy.

Please confirm that OpenAI will produce all documents reflecting policies, procedures, or practices related to the use of intellectual property in AI Models, including, *but not limited to*, Copyright Shield. That includes, for example, documents reflecting the recently reported change to ChatGPT's guidelines concerning compliance with copyright laws.[1]

4. <u>Narrowing of RAG-Related Requests (RFP Nos. 12, 13)</u>

RFPs 12 and 13 seek documents concerning "Defendants' use of Retrieval Augmented Generation (RAG)" and documents concerning "features or capabilities of [RAG]." In response, OpenAI states that it will produce "documents . . . sufficient to show how OpenAI uses RAG in connection with ChatGPT" (RFP 12) and "documents . . . sufficient to show OpenAI's understanding" of RAG (RFP 13).

These responses are insufficient. First, again, OpenAI does not state whether it is seeking to withhold responsive documents on the basis of its objections (and which ones). Please amend your responses accordingly. Second, the limitations that OpenAI appears to be making are improper. These requests are relevant to, among other issues, the Complaint's allegations about how Defendants' use of RAG infringes on The Times's copyrights and harms The Times and the substitutive effect of Defendants' products. Please confirm that OpenAI will produce documents concerning Defendants' use of RAG across the full scope of their Generative AI Products and Services rather than limiting this request to ChatGPT. Finally, OpenAI's "sufficient to show"

---

[1] *See* Ricky Sutton, ChatGPT Quietly Changes Its Mind on Copyrighted News Content, FUTURE MEDIA (Apr. 3, 2024), https://rickysutton.substack.com/p/chatgpt-quietly-changes-its-rules?utm_campaign=529baefa53-2024_04_04-incontext&utm_term=0_7a1e151592-529baefa53-399440900.

April 19, 2024
Page 4

approach to these Requests, which go to the heart of The Times' claims, is not appropriate. Please let us know whether you will reconsider your approach to RFPs 12 and 13.

   5.   Legislative or Executive Inquiries (RFP No. 2)

   The Times requested that OpenAI produce all "Documents You have gathered to submit or have submitted to any legislative or executive agency, committee, or other governmental entity in the United States that concern or relate to the allegations in the Complaint." Such documents would include, for example, documents OpenAI has gathered in response to the FTC's request for documents related to OpenAI's corporate structure, LLM-based products, and training and development of LLMs. OpenAI's categorical refusal to produce documents in response to this request is inappropriate, as those documents are plainly relevant and easy to collect given that they have already been gathered and produced to other entities. *See, e.g.*, Compl. ¶¶ 55–74 (discussing OpenAI's business model and collaboration with Microsoft); 83–101 (discussing the unauthorized reproduction of Times Works during training and in GPT models). Please let us know whether you will reconsider your approach to RFP No. 2.

   6.   Claims of Copyright Infringement (RFP No. 4)

   OpenAI limits and redefines The Times's RFP No. 4 regarding claims of copyright infringement by confining its response to documents sufficient to show "procedures governing *formal, written* complaints." R&Os at 10 (emphasis added). Please confirm that OpenAI will conduct a reasonable search for and produce all responsive, non-privileged documents regarding any claim of copyright infringement regardless of whether it is a "formal, written complaint." Please confirm that you will also produce those documents. The request also seeks documents concerning steps taken to respond to claims of infringement and to prevent infringement. Please confirm that your agreement to provide documents concerning "OpenAI's responses, if any, to such complaints" includes all steps taken to respond to claims and to prevent infringement, in addition to written responses sent in response to claims of infringement.

   7.   Alternatives to Using Copyrighted Content (RFP No. 6)

   RFP No. 6 seeks documents concerning alternatives to using copyrighted content to train the Text Generation AI Models without compensation. OpenAI's response is limited to "documents . . . *constituting* agreements with third parties relating to training data that was used to train the models used for ChatGPT." This is deficient for several reasons. First, the request is not limited to documents that constitute agreements but includes correspondence concerning those agreements, including negotiations with third parties whether or not they resulted in agreements. Second, it includes OpenAI's discussions (whether internal or external) concerning alternatives to using copyrighted content without compensation. Third, OpenAI's response entirely ignores the portion of the request related to the use of documents in the public domain. Please confirm that you will agree to produce these further categories of documents in response to this request.

April 19, 2024
Page 5

8. Impact of AI Models on Journalism (RFP No. 7)

RFP No. 7 seeks documents concerning the impact of AI Models on Journalism, as well as the use of AI Models as a source of Journalism, *or a substitute or product in competition with Journalism*. Your response is limited to documents sufficient to show "OpenAI's understanding of ChatGPT's effect, if any, on journalism." This response is unduly narrow to the extent that it does not address the portion of the request related to documents concerning the use of products as a source of, substitute for, or product in competition with Journalism. Moreover, the limitation to documents reflecting OpenAI's supposed "understanding" is unclear and improper; to the extent that OpenAI possesses documents that reflect the impact of AI Models on Journalism, those documents should be produced, regardless of whether those documents are consistent with OpenAI's "understanding" or not. Please confirm that you will agree to produce these further categories of documents in response to this request.

9. OpenAI's Transition to For-Profit Company (RFP No. 11)

RFP No. 11 seeks documents concerning OpenAI's transition to a for-profit company. OpenAI's response is limited to documents sufficient to show the for-profit status of each OpenAI Defendant and any changes to that status. That improperly excludes documents concerning the transition, including OpenAI's reasons for making changes to the for-profit status of each entity, which bears on the purposes and benefits of those entities' development of the at-issue products. Please confirm that you will agree to produce these further categories of documents in response to this request.

10. Collaborations Between Defendants Relating to the Generative AI Models (RFP No. 15)

OpenAI has offered to meet and confer regarding The Times's request for documents "concerning any commercial or supercomputing collaborations between and among the Defendants" in response to The Times's RFP No. 15. The Times accepts that invitation. The relationship between Microsoft and OpenAI is directly relevant to numerous claims and potential defenses in this case, and this request is not unduly burdensome or disproportionate to the needs of the case. Please confirm that OpenAI will produce documents responsive to this request. To ensure that the meet and confer is productive, please come prepared to discuss the types of documents in OpenAI's possession, custody, or control that are responsive to this request, including any documents that OpenAI contends are "unrelated to ChatGPT." R&Os at 20. As stated above, the Complaint implicates many products and services aside from ChatGPT. Compl. ¶¶ 16, 66–74, 108–23.

April 19, 2024
Page 6

Respectfully,

Zach Savage